

CLAIMS

1. A product of a process of integration of mixed format data, said product comprising at least one storage media device including a database of integrated data, said product produced by an integration process comprising the steps of:
 - accessing a database of structured data, the structured data comprising a set of data tuples;
 - accessing a source of unstructured data, the unstructured data including free text relatable to the data tuples of the structured data;
 - interpreting the free text to produce a set of construed data reflecting at least one relational fact conveyed in the free text, each construed datum relatable to a data tuple of the structured data;
 - integrating the produced data with the data tuples of the structured data; and
 - storing the integrated data produced by said integrating to the at least one storage media device in the form of a database.
2. A product of the process recited in claim 1, wherein said interpreting the free text includes the steps of:
 - identifying roles within the parsed text records, said identifying producing attribute extractions, each of said attribute extractions containing attribute information of the derived source text; and
 - applying caseframes to the attribute extractions, said applying caseframes producing a filtered set of attribute extractions.
3. A product of the process recited in claim 2 wherein said interpreting the free text further includes the step of identifying domains of the filtered set of attribute extractions.
4. A product of the process recited in claim 2, wherein said interpreting the free text further includes the step of assigning domain roles.
5. A product of the process recited in claim 4, wherein said interpreting the free text further includes the step of producing relation types.

6. A product of the process recited in claim 1, wherein:

said interpreting the free text further includes the step of creating a new database containing the structured data element produced in said producing; and

said product includes the created database.

7. A product of the process recited in claim 1, wherein:

said interpreting the free text further includes the step of producing a file containing the structured data element produced in said producing; and

said product includes the produced file.

8. A product of the process recited in claim 1, wherein the produced file has a format selected from the group of XML, character separated values, spreadsheet formats and file-based database structures.

9. A product of the process recited in claim 1, wherein said the stored integrated data includes reference information to the original free text for construed data.

10. A product of the process recited in claim 1, wherein:

the process further includes the step of constructing a library containing extracted attributes; and

said product includes the constructed library.

11. A product of the process recited in claim 1, wherein said interpreting the free text further includes the step of combining like relation types.

12. A product of the process recited in claim 1, wherein said interpreting the free text further includes the step of combining like attributes.

13. A product of the process recited in claim 1, wherein said interpreting the free text further includes the step of coalescing the identified attributes.

14. A product of the process recited in claim 1, wherein said interpreting the free text further includes the step of coalescing identified relation types.

15. A product of a process of integration of mixed format data, said product comprising a memory device having stored thereon integrated data, said product produced by an integration process comprising the steps of:

- accessing a database of structured data, the structured data comprising a set of data tuples;
- accessing a source of unstructured data, the unstructured data including free text relatable to the data tuples of the structured data;
- interpreting the free text to produce a set of construed data reflecting at least one relational fact conveyed in the free text, each construed datum relatable to a data tuple of the structured data;
- integrating the produced data with the data tuples of the structured data; and
- storing the integrated data produced by said integrating to the memory device.

16. A product of the process recited in claim 15, wherein said interpreting the free text includes the steps of:

- identifying attributes within the parsed text records, said identifying producing attribute extractions, each of said attribute extractions containing attribute information of the derived source text; and
- applying caseframes to the attribute extractions, said applying caseframes producing a filtered set of attribute extractions.

17. A product of the process recited in claim 16, wherein said interpreting the free text further includes the step of identifying domains of the filtered set of attribute extractions.

18. A product of the process recited in claim 17, wherein said interpreting the free text further includes the step of assigning domain roles.

19. A product of the process recited in claim 18, wherein said interpreting the free text further

includes the step of producing relational facts.

20. A product of the process recited in claim 15, wherein:

said interpreting the free text further includes the step of creating a new database containing the structured data element produced in said producing; and

said product includes the created database.

21. A product of the process recited in claim 15, wherein:

said interpreting the free text further includes the step of producing a file containing the structured data element produced in said producing; and

said product includes the produced file.

22. A product of the process recited in claim 15, wherein the produced file has a format selected from the group of XML, character separated values, spreadsheet formats and file-based database structures.

23. A product of the process recited in claim 15, wherein said the stored integrated data includes reference information to the original free text for construed data.

24. A product of the process recited in claim 15, wherein:

the process further includes the step of constructing a library containing extracted attributes; and

said product includes the constructed library.

25. A product of the process recited in claim 15, wherein said interpreting the free text further includes the step of combining like relation types.

26. A product of the process recited in claim 15, wherein said interpreting the free text further includes the step of combining like attributes.

27. A product of the process recited in claim 15, wherein said interpreting the free text further

includes the step of coalescing the identified attributes.

28. A product of the process recited in claim 15, wherein said interpreting the free text further includes the step of coalescing identified relation types.